Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm

Richard R. Irish, John L. Barker, Samuel N. Goward, and Terry Arvidson

Abstract

A scene-average automated cloud-cover assessment (ACCA) algorithm has been used for the Landsat-7 Enhanced Thematic Mapper Plus (ETM+) mission since its launch by NASA in 1999. ACCA assists in scheduling and confirming the acquisition of global "cloud-free" imagery for the U.S. archive. This paper documents the operational ACCA algorithm and validates its performance to a standard error of ±5 percent. Visual assessment of clouds in three-band browse imagery were used for comparison to the five-band ACCA scores from a stratified sample of 212 ETM+ 2001 scenes. This comparison of independent cloud-cover estimators produced a 1:1 correlation with no offset. The largest commission errors were at high altitudes or at low solar illumination where snow was misclassified as clouds. The largest omission errors were associated with undetected optically thin cirrus clouds over water. There were no statistically significant systematic errors in ACCA scores analyzed by latitude, seasonality, or solar elevation angle. Enhancements for additional spectral bands, per-pixel masks, land/water boundaries, topography, shadows, multidate and multi-sensor imagery were identified for possible use in future ACCA algorithms.

Introduction

A primary goal of the Landsat-7 (L7) mission is to populate the U.S.-held Landsat data archive with seasonally refreshed, essentially cloud-free Enhanced Thematic Mapper Plus (ETM+) imagery of the Earth's landmasses. To achieve this goal, the Landsat Project Science Office (LPSO) at NASA's Goddard Space Flight Center (GSFC) developed the Long-Term Acquisition Plan (LTAP): a mission-long imaging strategy designed to optimize the 250 scenes acquired each day by the ETM+ (Arvidson *et al.*, 2001, Arvidson *et al.*, 2006). An optimized scene acquisition has two primary characteristics: a priority for acquisition on that date and a low estimate of cloud contamination. A key element in the LTAP is a 12-month global analysis of vegetation derived from Advanced Very High Resolution Radiometer (AVHRR) observations using the Normalized Difference Vegetation Index (NDVI) (Goward *et al.*, 1999). Use of the resulting seasonality increases the probability of ETM+ collects during periods of heightened biological activity. Another key element of the LTAP strategy is to use cloud-cover (CC) predictions to reduce cloud contamination in acquired scenes.

In addition to the LTAP, acquisition scheduling by mission planners also requires reliable CC reports for imagery that is already acquired. Therefore, an automated cloud-cover assessment (ACCA) algorithm was created for determining the cloud component of each acquired ETM+ scene. The resulting CC assessment scores are used to monitor LTAP performance and reschedule acquisitions as necessary. The purpose of this paper is to document and evaluate the operational ACCA algorithm and to suggest potential enhancements for future Landsat-type missions.

Landsat-7 Mission Planning

To predict the probability of clouds in upcoming acquisitions, the L7 LTAP employs historical CC patterns developed by the International Satellite Cloud Climatology Project (ISCCP) and daily predictions provided by NOAA's National Centers for Environmental Prediction (NCEP). Candidate LTAP acquisitions are prioritized according to the forecasted cloud environment normalized against the historical CC average, as well as other system and resource constraints (Arvidson *et al.*, 2006). The priority for a candidate acquisition receives a boost if the forecasted CC is lower than the historical average (Gasch and Campana, 2000). The result of the scheduling process is an imaging schedule for the top 250 (on average) prioritized scenes. A schedule is transmitted to the satellite every 24 hours and forms the basis for operating the ETM+ during its 17 percent maximum daily duty cycle.

These 250 scenes, once acquired, are transmitted to the U.S. Geological Survey's Earth Resources Observation and Science (USGS/EROS) facility in Sioux Falls, South Dakota. The Landsat Processing System (LPS) processes the raw data into radiometrically uncalibrated and geometrically unresampled imagery; generates the associated browse imagery, ACCA scores, and other metadata; and sends the data set to the Landsat Archive Manager (LAM) for storage and eventual distribution.

Richard R. Irish is with SSAI, GSFC Code 614.4, Greenbelt, MD 20771 (rirish@pop400.gsfc.nasa.gov).

John L. Barker is with NASA/GSFC Code 614.4, Greenbelt, MD 20771.

Samuel N. Goward is with the Department of Geography, 2181 LeFrak Hall, University of Maryland, College Park, MD 20782.

Terry Arvidson is with Lockheed Martin, GSFC Code 614.4, Building 33, Room G313, Greenbelt, MD 20771.

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 10, October 2006, pp. 1179–1188.

0099-1112/06/7210–1179/$3.00/0

During normal operations, the Landsat Mission Operations Center (MOC) located at GSFC receives the previous day's ACCA scores from the LAM. The scheduling software uses the scores to separate successful acquisitions (those with sufficiently low CC) from those that require re-imaging due to high CC. Routine comparisons of ACCA scores to the CC predictions are made to evaluate the reliability of the forecast information; these comparisons treat the ACCA scores as "truth." The accuracy of the ACCA process is essential to the efficient LTAP refreshing of the global archive with cloud-free imagery.

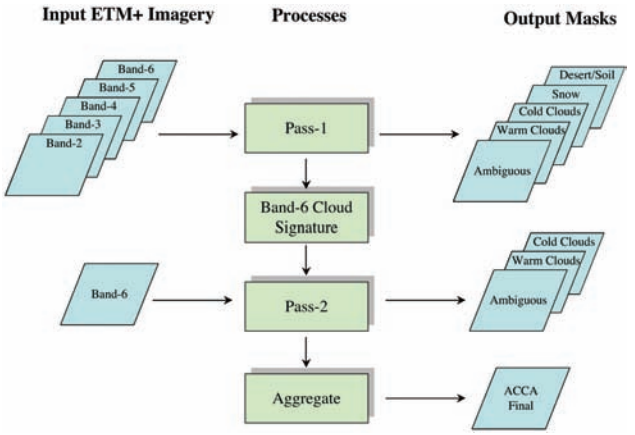

Plate 1. Overview of L7 ETM+ automated cloud-cover assessment (ACCA) algorithm software flow.

Automated Cloud-Cover Assessment (ACCA) Algorithm

Many of the essential elements of the ETM+ ACCA algorithm have been previously described together with the heritage browse-based ACCA algorithm used in processing Thematic Mapper (TM) imagery from the Landsat-4 (L4) and Landsat-5 (L5) satellites and the computer-driven limitations of both (Irish, 2000).

The L7 ACCA algorithm is an unsupervised classifier for clouds, which takes advantage of known spectral properties of clouds, snow, bright soil, vegetation, and water. The primary goal of the algorithm is to quickly produce acceptable scene-average estimates for CC during initial LPS processing. It was not intended to produce a "per-pixel" mask indicating the presence or absence of clouds for every pixel in L7 imagery. L7 "ACCA clouds" are defined as optically thick or nearly opaque because the ETM+ spectral bands do not easily detect semi-transparent clouds such as Cirrus Uncinus (i.e., "mare's tail"), Cirrus Fibratus, and cloud edges. Shadows from clouds are also not assessed. Furthermore, if all cirrus clouds were detected and used as a criterion to "reject" scene acquisitions, then most acquisitions would be "rejected" because of the pervasive character of thin cirrus clouds in the majority of the 183 km by 180 km L7 scenes.

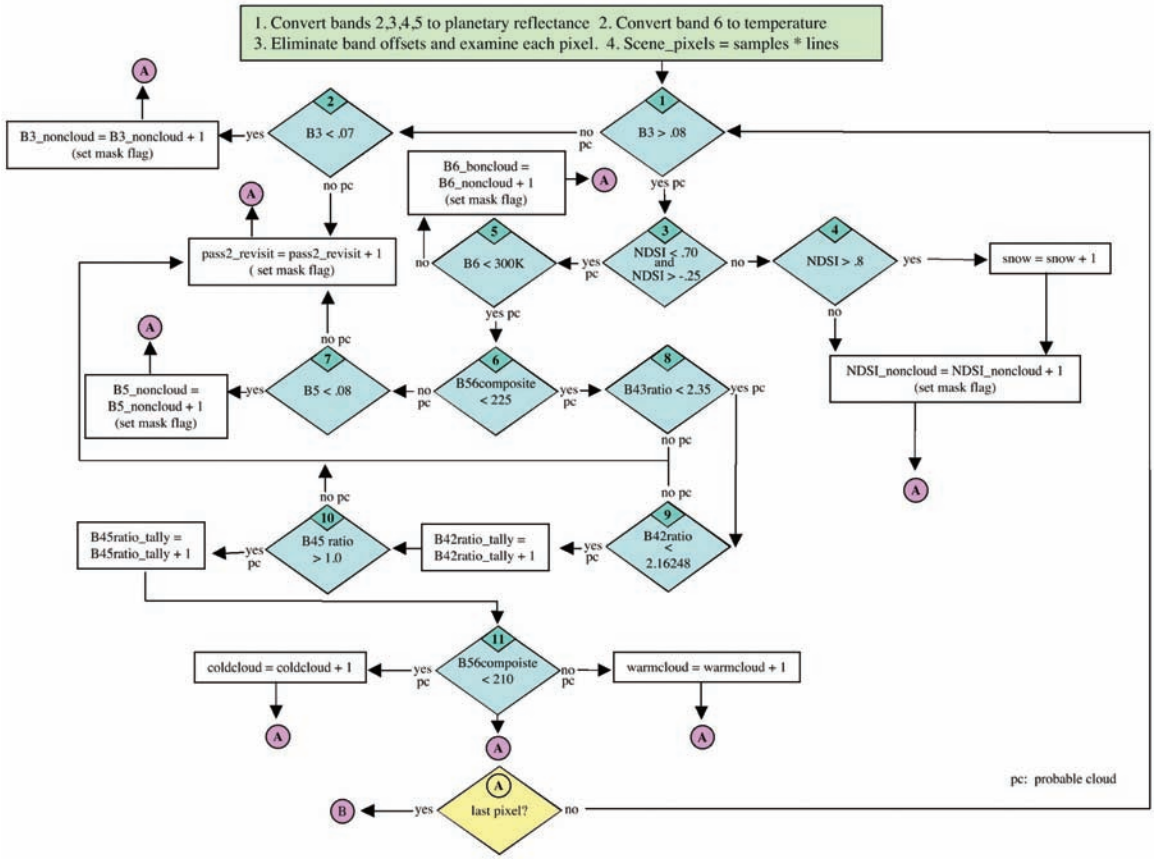

Plate 2. ETM+ ACCA flow chart (1 of 4): spectral cloud identification.

Elements of ACCA Algorithm

ACCA uses five of the eight ETM+ spectral bands:

- Band-2 (B2): 0.53 to 0.61 μm, Green, 30 m resolution
- Band-3 (B3): 0.63 to 0.69 μm, Red, 30 m resolution
- Band-4 (B4): 0.78 to 0.90 μm, Near Infrared, 30 m resolution
- Band-5 (B5): 1.55 to 1.75 μm, Shortwave Infrared, 30 m resolution
- Band-6 (B6): 10.4 to 12.5 μm, Thermal Infrared, 60 m resolution.

Bands 2 through 5 are converted to planetary spectral reflectance (Markham and Barker, 1986), and Band-6, in its low gain form, is converted to an at-satellite apparent brightness temperature (Kelvin) as per the L7 Science Data Users Handbook (Irish, 1999).

The ACCA algorithm consists of twenty-six specific decisions or filters. Plate 1 illustrates the processing overview. The algorithm flow is divided into four processes:

1. *Pass-1 Spectral Cloud Identification* (Plate 2) where the five input images for B2 through B6 are used to identify warm-cloud, cold-cloud, possible-cloud (ambiguous), snow, and non-cloud masks.
2. *B6 Cloud Signature Development* (Plate 3) from the known cloud pixels.
3. *Pass-2 Thermal Band Cloud Separation* (Plate 4), performed by re-examining ambiguous pixels using only B6. Those that qualify are assigned to either a warm or cold cloud class.
4. *Image-Based CC Assignments, Aggregation and Filling* (Plate 5) aggregates Pass-1 and Pass-2 clouds and fills mask cloud holes using nearest-neighbor resampling.

The initial ETM+ ACCA algorithm was developed prior to the L7 launch in 1999 from analysis of about 500 L5 TM scenes. During the three-month On-Orbit Initialization and Verification Period (OIVP) for ETM+, it was necessary to change the values of the ACCA parameters before satisfactory CC estimates were obtained. Only minor adjustments to the parameters have been made since ETM+ began acquiring operational imagery 29 June 1999. Complete algorithm details can be found in the L7 Science Data Users Handbook.

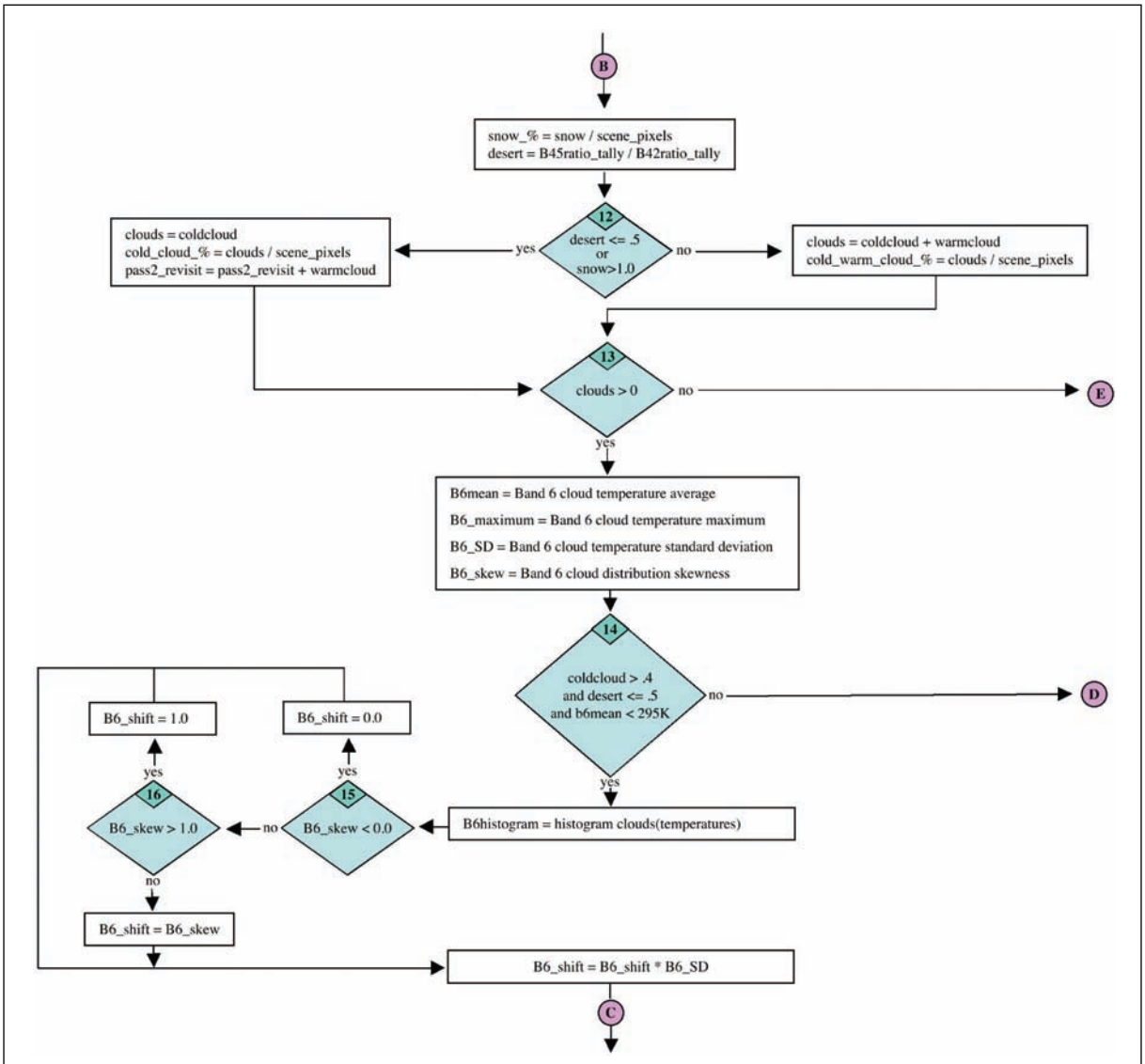

Plate 3. ETM+ ACCA flow chart (2 of 4): band 6 (B6) cloud signature development.

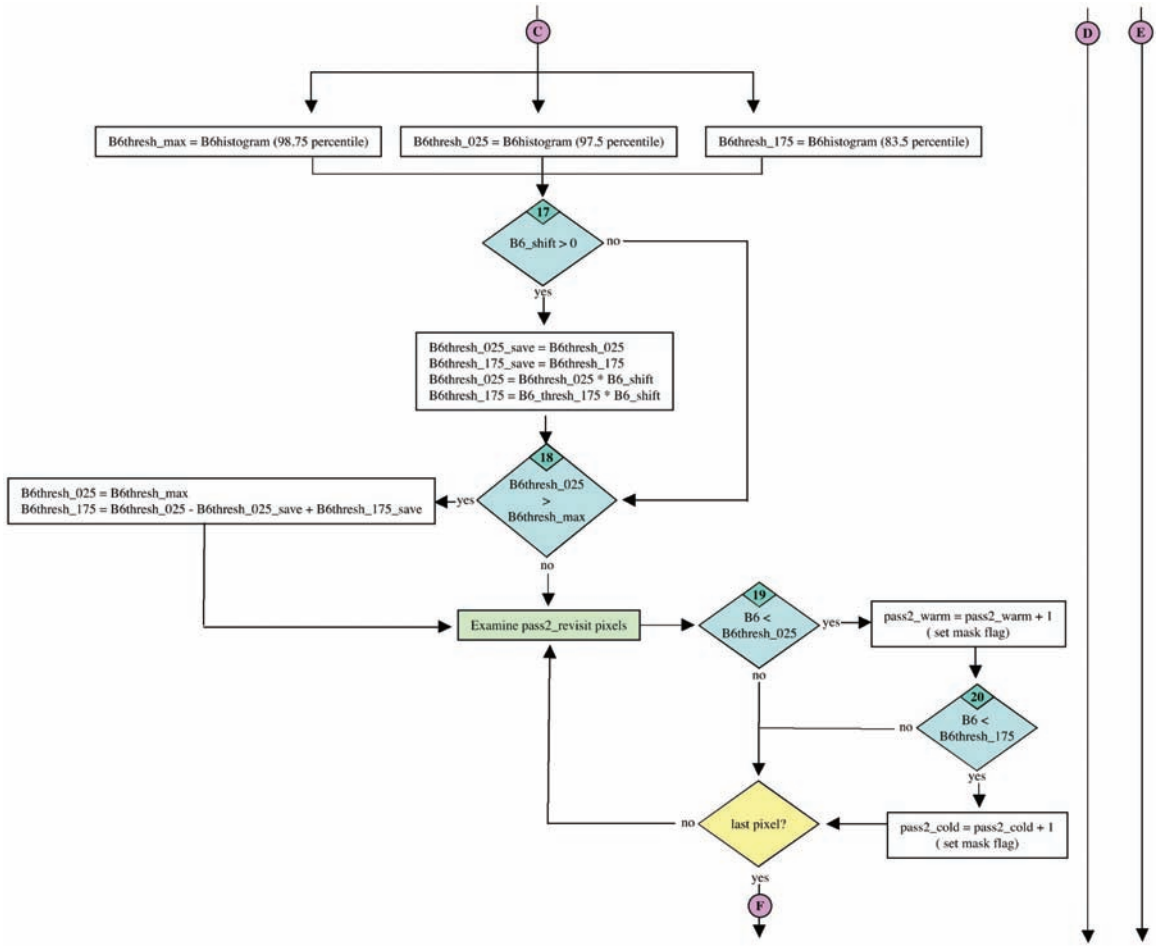

Plate 4. ETM+ ACCA flow chart (3 of 4): pass-2 thermal band cloud separation.

Validation of ACCA Algorithm

Following the first few years of L7 operations, the LPSO staff undertook a validation of all elements of the LTAP including ACCA (see Arvidson *et al.*, 2006; Markham *et al.*, 2006). The LTAP was essentially stable by 2001, therefore the approximately 83,000 ETM+ scenes in the U.S. archive for calendar year 2001 were sampled to evaluate ACCA performance. The approach chosen to validate ACCA performance was supervised classification of a stratified sample using the LPS-produced three-band browse imagery for comparison to corresponding ACCA scores.

LPS browse images are produced by wavelet compression and subsampling of radiometrically-corrected imagery by a factor of 64 (8×8). Browse imagery at a spatial resolution of 240 meters (**http://glovis.usgs.gov/**) is produced. B5, B4, and B3 are used to define the red (R), green (G), and blue (B) elements respectively. A linear contrast stretch is applied to each browse image to ensure adequate contrast in the RGB product. A reduced-resolution browse-scene approach to assessing CC was feasible because ACCA represents a single number for the entire scene. Furthermore, the ACCA definition of clouds as nearly opaque and the ability to visually compare the supervised classification of browse imagery to adjacent scenes and other dates allowed an iterative and precise analysis approach. This visual cloud-cover assessment (VCCA) was used as a measure of the true CC in the scene.

Visual Cloud-Cover Assessment (VCCA) Validation Procedure

Browse imagery was downloaded to a desktop computer as 24-bit RGB files. Calculations were accomplished using Adobe® Photoshop® image processing software with 24-bit precision. Edges of the 825-column by 750-line browse input imagery, where the bands are offset from each other, were first trimmed to a common size of 775 columns by 750 lines. The *magic wand* and *freehand lasso* tools of Photoshop® were subsequently used to isolate clouds. The wand employs a seed-fill threshold algorithm to compute regions of brightness-similarity based on a mouse click of a single pixel. The algorithm compares the selected pixel's brightness values to all other pixels and retains those within a selectable tolerance threshold. For example, clicking on an RGB-browse-image pixel with values R:200, G:220, and B:240 and a tolerance set to 5, results in selection of pixels in the ranges: $195<R<205$, $215<G<225$, and $235<B<245$. Additional cloud pixels were added by using the wand repeatedly until the cumulative selection of visible clouds had essentially zero possibility of VCCA omission errors. Snowfields and other unwanted bright features were then manually subtracted using the lasso tool to reduce VCCA commission errors. After the VCCA scores were established, the resulting cloud pixels were filled with 255s and all others a value of zero. The result was a binary cloud mask that allowed a CC percentage computation (Plate 6) that served as the cloud "truth" for validating the accuracy of the official ACCA USGS/EROS scores.

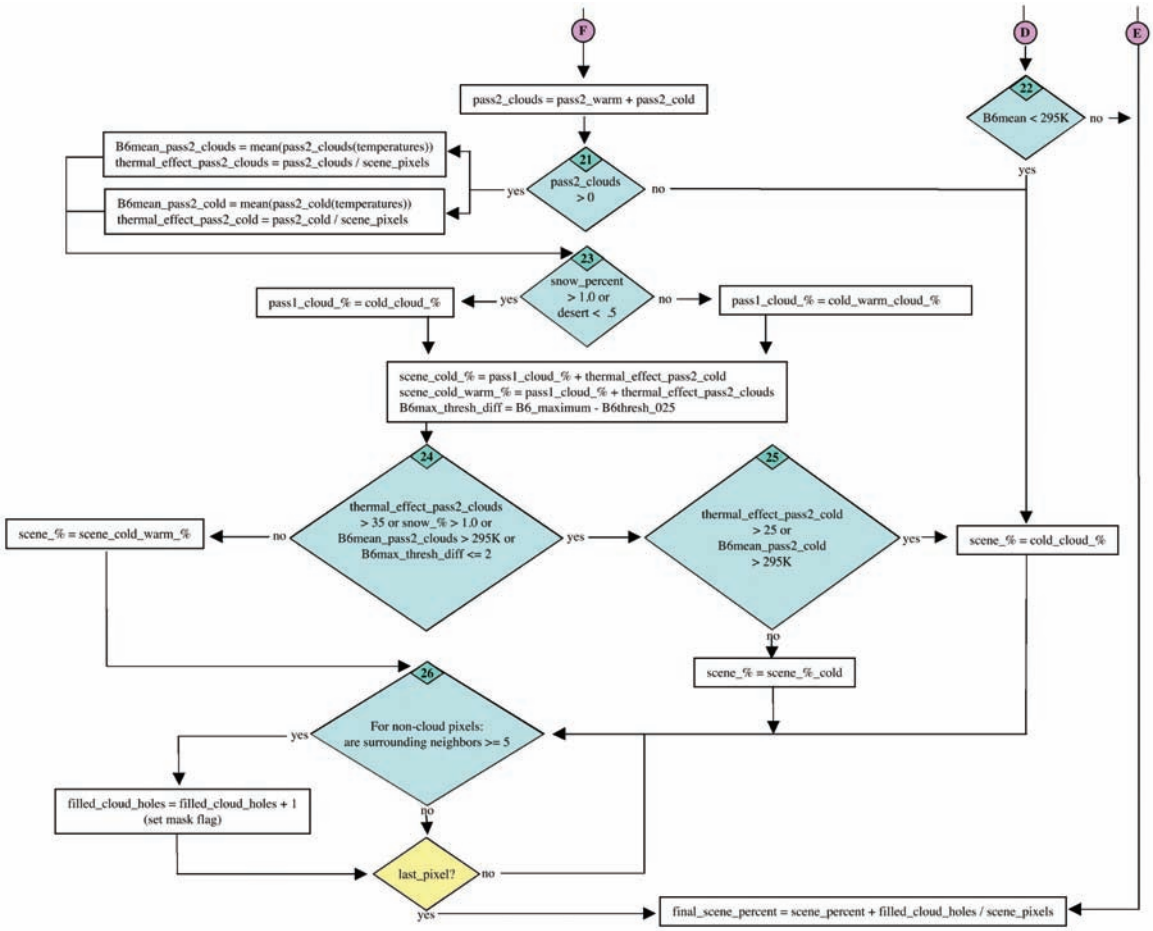

Plate 5. ETM+ ACCA flow chart (4 of 4): image-based assignments, aggregation, and filling.

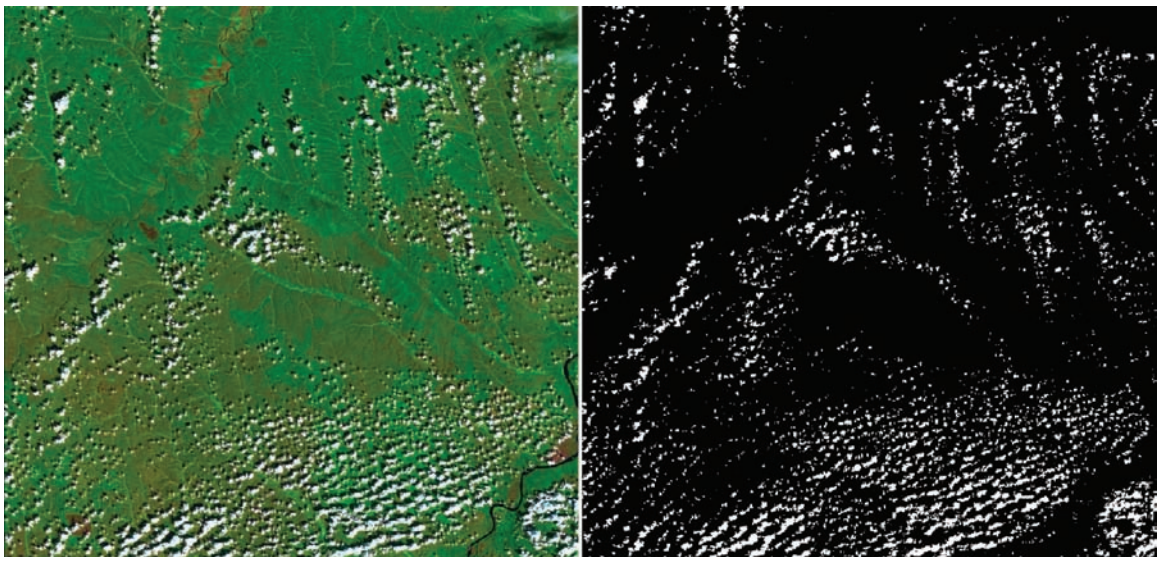

Plate 6. Illustrative ETM+ B3/B4/B5 RGB USGS/EROS browse image and Photoshop® VCCA (visual cloud-cover assessment) mask of "true" cloud-cover (CC).

Systematic Random Sample Design

The Earth was systematically subdivided into nine latitudinal zones on the basis of environmental equivalence, independent of the amount of land area within each zone. The six mid-latitude zones span 15° while the Equatorial and Polar zones cover 30°. For eight of these zones, 21 Worldwide Reference System (WRS) locations were randomly selected from the approximate 4,000 unique daylight WRS locations available over land. For the Polar South zone, only 20 samples were selected (the reason for this is explained below). If a selected WRS location was dominated by water, the sample was replaced. The resulting 188 WRS locations are illustrated in Plate 7.

For each of the 188 randomly selected WRS locations, the "peak NDVI" scene was selected from the time interval when the solar zenith angle was within 10° of minimum at that location. This produced a systematic random sample design focused on the vegetative growth peak during the growing season.

Finally, for each zone except the Polar South, one of the 21 WRS locations was randomly selected randomly for analysis of seasonal variability. For this one WRS location per zone, additional scenes were selected from the 2001 U.S. archive for the three remaining seasons of the year, namely winter, spring, and fall. The peak-NDVI scenes from these eight seasonal sets were removed from the original 188 locations leaving 180 scenes for the ACCA/VCCA peak-NDVI analysis. This was done to ensure unique seasonal analysis. No seasonal study was attempted for the Polar South zone due to the limited imaging window for Antarctica.

These peak-NDVI and seasonal samplings were designed to test for possible ACCA-dependence on latitude, solar illumination angle, and season. The histogram of ACCA scores from this sampling was statistically identical to that for the 83,000 2001 scenes, which provided an independent validation of the sampling randomness.

ACCA Validation Results

The first concern that was addressed was the possibility that errors were induced in the VCCA process due to the difference in processing between browse imagery and full-resolution imagery. The browse-reduction process conceals clouds less than about 240 meters in size, resulting in possible VCCA omission errors. ACCA examines each pixel at nominal resolution. A systematic VCCA omission error was not observed, which eliminated this concern.

ACCA Versus VCCA Validation

The 212 ACCA scores for the 180 peak-NDVI and 32 seasonal scenes were retrieved from the USGS/EROS LAM and compared to the browse truth scores obtained by the manual VCCA procedure described above. A validation curve of automated ACCA versus VCCA scores for the 180 high-NDVI scenes is given in Plate 8. A perfect validation would correspond to a linear regression fit with an intercept through the origin, a slope of unity and zero dispersion from the fit line. This would validate both the ACCA and VCCA scores.

The high-NDVI sample had three statistically identifiable outliers. The three outliers in the random sample equate to

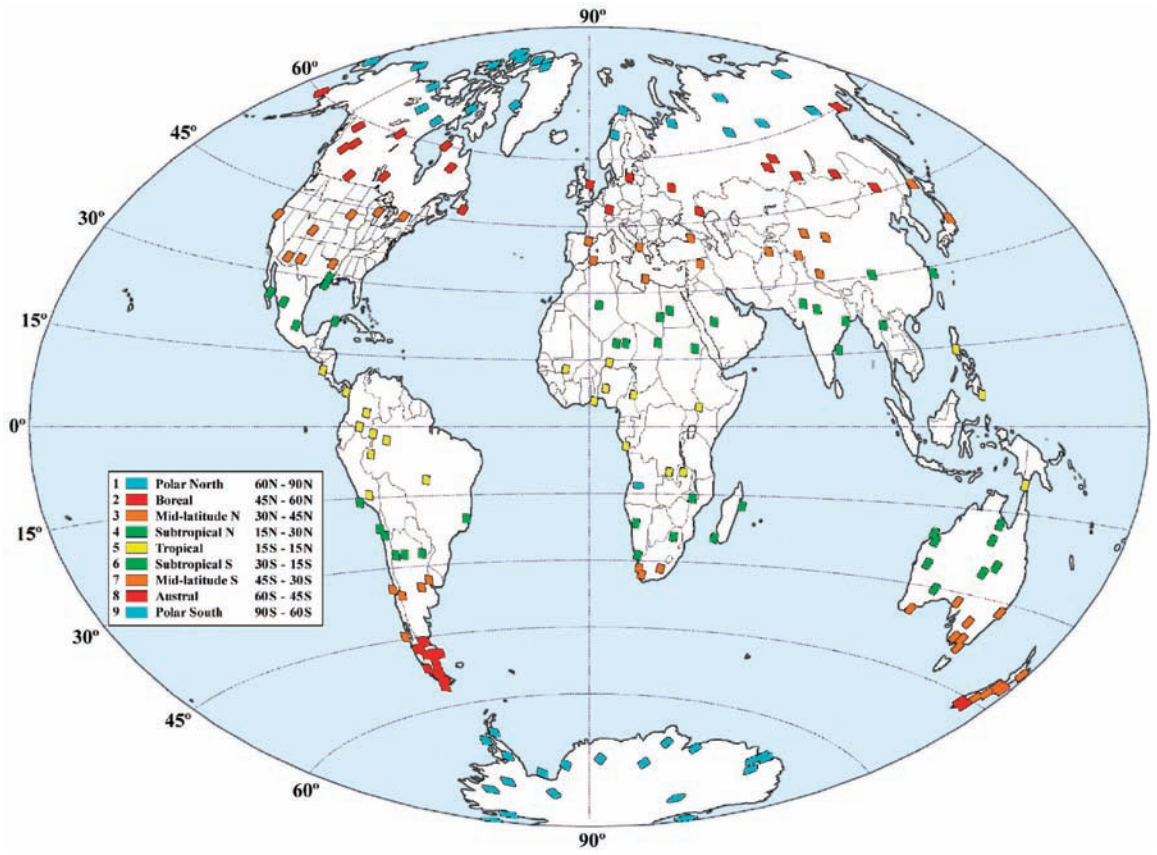

Plate 7. 188 random Landsat WRS locations in nine latitudinal zones.

an expectation that ACCA will under-perform by more than ±15 percent at a 2 percent rate. Errors of omission and commission are discussed in the outliers section. The three outliers were omitted from the regression fit and are labeled as 1, 2, and 3 in Plate 8.

The observed fit between the remaining 177 scenes allows for acceptance of the null hypothesis that the intercept is zero and the slope is unity within a standard error of ±5 percent. The variance unaccounted for by the linear fit is 2 percent. The residuals are random. Consequently, the expected precision of the ACCA algorithm is ±5 percent in 98 percent of all 2001 scenes.

Four of the 180 high-NDVI images, or about 2 percent, contained electronic band-specific gain changes within a scene from intentional commands to change the gain state. Currently, ETM+ ACCA images are radiometrically corrected in their entirety based on the gain state of the first scan line in a scene. Gain changes are not recognized during ACCA processing which may engender discordant CC scores related to individual bands, location in a scene, and imaged ground features. The random scene sampling allowed a 2 percent sample of gain change scenes to be representative of the 2001 archive. Therefore, the gain-change scenes were retained. None exhibited a statistically significant deviation from the regression line. However, a sample of four is too few to validate the scene-average ACCA for scenes with gain changes. Tracking and adjusting for gain changes is necessary if accurate per-pixel cloud masks are generated.

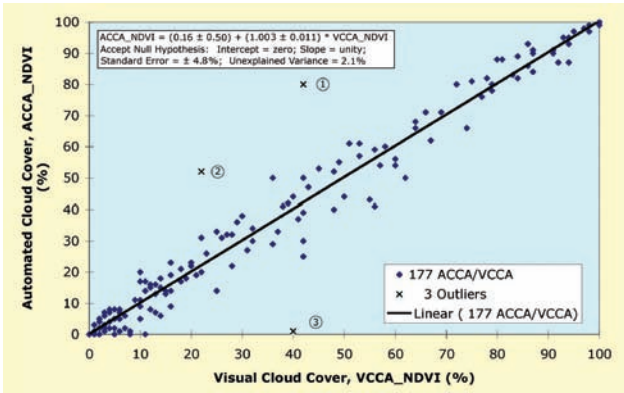

Plate 8. Peak-NDVI ACCA/VCCA validation curve (±5 percent) for 180 WRS locations in 2001.

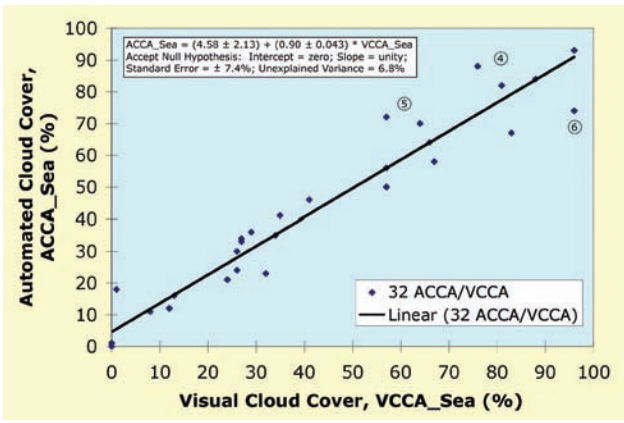

Plate 9. Four season ACCA/VCCA validation curve (±7 percent) for eight Landsat WRS locations in 2001.

ACCA Seasonal Independence

The results for non-peak NDVI are presented in Plate 9 for a sample of four 2001 seasonal scenes across eight WRS sites. The residuals from a linear fit were examined and no outliers were rejected. Three residuals are labeled as 4, 5, and 6 for later discussion of possible error sources. Two of the 32 seasonal scenes had gain changes, but no removal rationale existed. One gain-change scene, labeled 6 in Plate 9, had a large residual that is examined further in the outlier section.

The observed ACCA versus VCCA fit of the 32 seasonal scenes allows for acceptance of the null hypothesis that a zero intercept exists and the slope is unity but with a higher standard error of ±7 percent compared to ±5 percent for the peak-NDVI scenes. Variance unaccounted for using the linear fit is 7 percent. The seasonal residuals are random. Again, no evidence exists regarding a systematic bias in the ACCA scores. One can conclude that ACCA is an unbiased estimator of CC for seasonal data at a precision of ±7 percent.

ACCA Latitude Independence

One method to test for possible effects of latitude is to examine differences between ACCA versus VCCA validation curves for each of the nine stratified zones. The results of nine linear regressions with three outliers removed are summarized in Table 1. Within statistical uncertainty, each zone has a zero intercept and a slope of unity, which represents the criteria for an unbiased estimate of ACCA CC. Standard errors range from ±3 percent to ±5 percent with average unaccounted-for variance of 2 percent, which is consistent with the values from the collective set of 177 peak-NDVI scenes. The residuals from each zone are random. Collectively, these independent statistical assessments of ACCA validation for each zone are justification for stating that ACCA accuracy is latitude independent.

The 20 random samples drawn in the Polar South matched the ACCA standard error of ±5 percent in other zones. Other studies (Choi and Bindschadler, 2002) have demonstrated some enhancement in cloud detection over ice.

The conclusion is reached that ACCA is an unbiased estimator of CC for data taken at different latitudes to a precision of ±5 percent.

TABLE 1. REGRESSION RESULTS FOR LATITUDE-DEPENDENT ANALYSIS DEMONSTRATING NO SYSTEMATIC ACCA ERROR. FOR EACH ZONE THE NULL HYPOTHESIS OF ZERO INTERCEPT AND UNITY SLOPE WAS ACCEPTED

Zone	Samples	Intercept	SD	Slope	SD	SE	100 (1-RSQ)
1	20	1.079	0.866	0.978	0.019	2.93	0.75
2	20	0.956	1.944	0.989	0.040	5.97	2.96
3	19	1.785	1.577	0.990	0.039	5.13	2.78
4	20	0.664	0.705	1.046	0.025	2.61	1.06
5	19	−1.009	1.964	1.055	0.033	4.90	1.74
6	20	−1.461	0.775	0.936	0.036	2.93	2.71
7	19	0.864	1.422	1.018	0.037	4.71	2.29
8	20	3.721	2.220	0.952	0.037	4.87	2.73
9	20	−3.125	1.821	0.988	0.038	5.38	2.67
Average	19.667	0.386	1.477	0.995	0.034	**4.381**	**2.191**
SD	0.500	1.996	0.570	0.039	0.007	1.227	0.815

Solar Illumination Angle Independence

A cosine of the solar zenith angle was used to convert at-satellite radiances into planetary reflectance for use in ACCA, therefore no dependence on solar illumination angle was expected from analog data of infinite precision, based on radiance considerations alone. However, ETM+ is only an eight-bit radiometer and lower solar irradiance increases the percent uncertainty of lower at-satellite radiances due to limiting ETM+ signal-to-noise values. Consequently, a linear test for correlation of the ACCA error with solar illumination angle was performed. ETM+ acquisitions for the U.S. archive acquired after 2000 had no collects below 15° solar elevation, so ACCA evaluation below this angle was not possible. Comparison of ACCA/VCCA differences versus solar elevation angles revealed no systematic trend. The null hypothesis that the slope was unity and intercept zero within the uncertainty of a linear fit was accepted. A solar angle correlation was not observed in the data.

Characterization of ACCA Outliers and Extreme Residuals

Six unusual ACCA/VCCA pairs were identified and are highlighted in Plates 8 and 9. Re-examination of the VCCA images and the ACCA scores verified that five of these pairs represented incorrect ACCA values. One of the six, labeled as 5 in Plate 9, was actually a VCCA error where the ACCA algorithm had successfully detected transparent cirrus more thoroughly than VCCA. ACCA error analysis did not include this outlier. Primary ACCA commission and omission CC errors are summarized for the remaining five pairs in Table 2.

The largest commission errors are at high altitudes over scenes that have snow during summer. Mountainous snowfields can saturate B2 in high- and low-gain mode, causing snow to be classified as cloud during Pass-1. The result is a distorted pass-1 thermal profile, which leads to high commission errors. Though saturated B2 imagery occurs over snow, the ACCA algorithm could perform correctly by adjusting cloud-discerning parameters according to the elevation of the imaged terrain.

The largest omission errors in ACCA are associated with optically-thin cirrus clouds over water. The ACCA algorithm typically underestimates semi-transparent cirrus over water by about 30 percent. The lack of an ETM+ 1.33 or 1.88 μm cirrus-detection band renders clouds of this type invisible unless they are nearly opaque or exhibit a particularly cold thermal signal. In general, the B6 thermal responses for optically-thin cirrus are weak and contribute little to their recognition. The human eye visualizes subtle spatially-transparent cirrus cloud shapes in three-band browse over water, whereas the automated single-band algorithms within ACCA do not. The opposite is sometimes true over land as indicated by the VCCA error on pair 5 in Plate 9. A possible solution, discussed later, involves masking the oceans with a land/water mask.

Discussion: Proposed Enhancements

The operational L7 ACCA algorithm was developed with certain constraints that no longer apply because of today's increased computer processing and storage capabilities. As a result of seven years of on-orbit experience with the operational ACCA algorithm and the validation experience, it is now reasonable to suggest additional enhancements that represent added value for today's users and for future systems. Potential enhancements include using additional spectral bands, binary cloud masks, land/water boundary maps, elevation models, shadow masks, and multi-date and multi-sensor imagery.

Binary Cloud Mask Enhancement

We assume that cloud products for future 10 to 100 meter moderate spatial resolution sensors will define various types of clouds and cloud shadows as binary masks, which will accompany user-ordered products. For example, multi-date image compositing techniques are frequently used to generate cloud-free composites for direct radiometric modeling of both intensive and extensive variables. Other disciplines require same-day imagery for every pixel even if covered by thin cirrus clouds. A per-pixel cloud mask would support both types of users. The current L7 ETM+ ACCA algorithm internally generates masks that disappear after processing. These or similar ones should be saved and made available. Analogous binary cloud masks currently exist for other coarse-resolution sensors such as the Moderate-resolution Imaging Spectroradiometer (MODIS) (Platnick *et al.*, 2003).

Cirrus Spectral Band Enhancement

A portion of the existing "cloud-free" ETM+ archive, with an ACCA score of 10 percent or less, contains transparent cirrus clouds which the current ACCA algorithm cannot reliably identify. The specifications for the future Landsat Data Continuity Mission (LDCM) sensor call for spectral observations (1.38 μm band or equivalent) to identify cirrus clouds, similar to the EOS MODIS sensors on the Terra and Aqua satellites (Gao and Kaufman, 1995; Gao *et al.*, 1998; Brill, 2006). When a cirrus spectral band does become available for Landsat-like sensors, a new definition of non-opaque clouds will be required to indicate if at least some of the ground is visible through a cloud.

Panchromatic Spectral Band Enhancement

The higher spatial resolution 15 m B8 panchromatic pixels were not available for use in the current ACCA algorithm. When combined with internally-generated date-specific masks for commission classes such as snow, desert, and even water, B8 imagery affords a method for building cloud mask estimates that are more inclusive of cloud edges.

TABLE 2. CAUSE OF ACCA OMISSION AND COMMISSION ERRORS IN CLOUD COVER (CC) ESTIMATES FOR EXTREME ACCA/VCCA PAIR RESIDUALS FROM THE VALIDATION CURVE

Extreme Value	Plate	ACCA Score	VCCA Score	CC error	WRS Path	WRS Row	Date	Location	Cause	Fix
Outlier 1	8	80	42	+38	143	038	05/31/01	Tibet	Snow as Clouds; B2 Saturation	elevation
Outlier 2	8	52	22	+30	192	056	11/14/01	75% Ocean off Africa	Cold water as Clouds	land/water mask
Outlier 3	8	1	40	−39	073	090	12/14/01	Ocean off New Zealand	Cirrus over Water	land/water mask
Outlier 4	9	67	83	+16	231	093	08/11/01	Mountains in Chile/Argentina	Snow as Clouds	elevation
Outlier 5	9	74	96	−22	091	087	03/15/01	67% Ocean off Australia	Cirrus over Water	land/water mask

Multi-Spectral Band Enhancement

The current ACCA algorithm was limited to a single band or single variable thresholding. A more powerful approach to unsupervised classification is to use maximum likelihood classifiers on all the bands or on a transformed combination of bands.

Binary Land/Water Mask Enhancement

Currently, the ACCA algorithm computes CC scores for an entire scene regardless of the Earth features imaged, including large areas of ocean that interface with land. Mission operations problems arise when identified clouds exist only over water. Unacceptable cloud scores are reported, causing the MOC to reschedule acquisitions for scenes that meet the success criteria as observed in Plate 10. This scene would be slated for re-acquisition due to an unacceptable 34 percent ACCA score when in fact the land area was cloud-free. For some areas of the Earth, such as the west coast of Chile, this phenomenon is common, resulting in wasteful consumption of satellite acquisition time.

A prototype ACCA enhancement has been developed to apply an ocean-land mask to each ETM+ scene prior to cloud assessment. The mask employed is the 1 km sea-land mask developed as part of the Global Land AVHRR Data Set Project (Eidenshink and Faundeen, 1994). Separate ACCA parameters could also be used with the ocean-land to reduce the most significant omission errors of 30 percent (±10 percent) for thin cirrus over water, as reported in Table 2. Additionally, cloud-free imagery or mosaics can be used to validate and update the land/water mask's geometric precision.

DEM Topographic Enhancement

As noted previously, separation of clouds and snow under low sun illumination angles and at high altitudes is a problem that might be resolved with elevation information. The snow and cloud confusion occurs at higher elevations with regularity in Mongolia and Argentina during the winter months. Global digital elevation models (DEMs) exist that have a 1 km spatial resolution, sufficient for ACCA elevation-specific parameter adjustments. Dynamic scene-dependent parameters will reduce commission errors identifying snow as cloud regardless of solar illumination angle, including the two illustrated in Table 2.

Binary Cloud Shadow Mask Enhancement

The ACCA algorithm was designed to discern clouds but not their projected shadows. Clouds and cloud shadows measured together represent an enhanced indicator of total image usability. Three cloud-shadow analysis approaches have been identified. A spectrally-based approach assumes that darker pixels are shadowed surfaces and tests that hypothesis against a cloud mask for that image. A spectral approach is the most independent to the cloud masking algorithm, which potentially could lead to the highest omission errors. A second approach is temperature-based, where the B6 brightness temperature of the clouds in the cloud mask is used to estimate their altitude. Cloud altitude, sun azimuth, and sun elevation are used to calculate the expected cloud shadows. Expected errors are largely dependent on the estimated cloud altitude accuracies. A third minimum-omission cloud-shadow approach assumes a maximum height or an estimate based on a 2 to 3 standard deviation upper limit on the cloud altitude. An exaggerated shadow mask, large enough to ensure that pixels outside of it are shadow-free, is then projected.

Multi-Date Cloud Masking Enhancement

The power of viewing the same scene on different dates to identify clouds on any one date was realized during VCCA analysis. An automated multi-date compositing process is feasible because cloud distribution patterns vary from scene to scene. A comparison of imagery on dates before and after a scene of interest reduces potential errors of commission caused by land features such as bright sand and snow.

Multi-sensor Cloud Masking

The biggest advancement in CC identification may one day come from using cloud heights and cloud masks from coarser resolution sensor systems such as MODIS and Geostationary Operational Environmental Satellite (GOES). Assum-

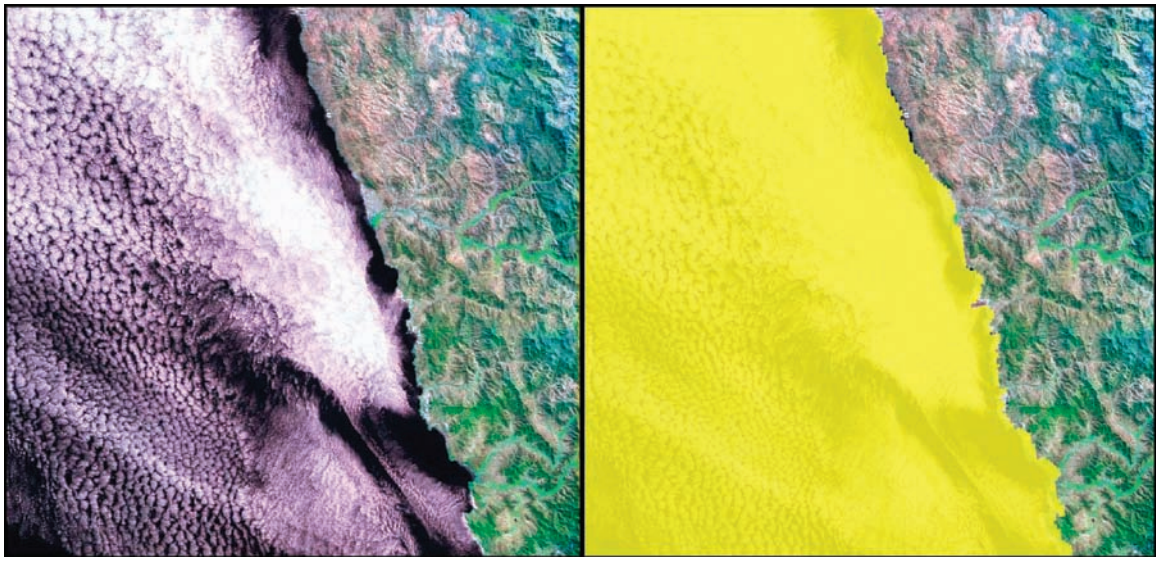

Plate 10. Illustrative current rejected image (ACCA = 34 percent CC) and proposed enhancement of accepted land-masked image (ACCA = 0 percent CC).

TABLE 3. STRENGTHS AND WEAKNESSES OF THE L7 ETM+ ACCA ALGORITHM

Strengths	Weaknesses
Fully automated methodology	No per-pixel cloud masks (only scene and quadrant scores),
Five of eight ETM+ spectral bands	
No ancillary information required	No cloud shadows mask
No geometric processing	Thin cirrus clouds often missed
Near-real time availability for browsing and scheduling	Ground fog occasionally missed (too warm)
No dependence on latitude, season, or solar zenith angle	Snow occasionally identified as cloud
Precise to ±5% for scene-averaged cloud-cover	

ing automated registration of data from multiple sources, the potential exists for interpolating cloud movement from half-hourly GOES images and/or employing the more information-rich cloud products from MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS). When a multi-sensor cloud-masking enhancement is available, elimination of CC errors of both omission and commission at aggregated pixel levels from moderate resolution imagery becomes possible.

Conclusions

The ACCA algorithm was designed to determine the quantity of cloud-contaminated pixels over land on an ETM+ scene basis. The pre-launch reliability goal for the ACCA algorithm was an accuracy of 10 percent actual CC for 95 percent of all LTAP imagery. The algorithm passes that reliability test for peak-NDVI acquisitions and narrowly misses for seasonal acquisitions.

The global Landsat image archive is the most relevant moderate spatial resolution (10 to 100 m) satellite land remote sensing archive in existence. From 2000 through 2005, USGS/ EROS-archived ACCA cloud-cover scores of 10 percent or less increased from 35 percent to 42 percent, while scores for cloud-cover greater than 50 percent correspondingly decreased from 33 percent to 28 percent. The assumption that no real change in global CC over land occurred during this period leads to the conclusion that the feedback of ACCA scores into the scheduling process has contributed to a 20 percent improved efficiency in acquiring cloud-free ETM+ imagery. As of 2006, more than 1.8 million Landsat scenes have been accumulated in the U.S. archive, including more than 600,000 scenes from ETM+. The percentage of usable cloud-free or low-cloud-content imagery is significantly higher in the ETM+ archive, as compared to the other sensors.

This ACCA validation documents the strengths and limitations of the operational scene-average L7 ACCA algorithm (Table 3). It may also serve as a departure point for ACCA approaches to be used in future Landsat-type missions.

The suggested ACCA enhancements are important in a century that requires cloud-free mosaics of the Earth for monitoring and modeling environmental change. Future Earth remote sensing systems will place increasing importance on the ability to remove imaged clouds. Multi-date compositing facilitated by increased revisits and multiple satellites perhaps represents the best approach. Until then, a finely-tuned, enhanced ACCA algorithm represents an important tool for research scientists to study the biosphere with an unobstructed view.

References

Arvidson, T., J. Gasch, and S.N. Goward, 2001. Landsat 7's long term acquisition plan – An innovative approach to building a global archive, Special Issue on Landsat 7, *Remote Sensing of Environment*, 78:13–26.

Arvidson, T., S.N. Goward, D.L. Williams, and J. Gasch, 2006. Landsat-7 long-term acquisition plan: Development and validation, *Photogrammetric Engineering & Remote Sensing*, 72(10).

Brill, J., 2006. Landsat Data Continuity Mission, Goddard Space Flight Center, URL: http://ldcm.nasa.gov/index.htm (last date accessed: 22 July 2006).

Choi, H., and R. Bindschadler, 2002. *Cloud Detection in Landsat Imagery of Ice Sheets Using Shadow Matching Techniques and Automatic Normalized Difference Snow Index Threshold Value Decisions*, Goddard Space Flight Center White Paper.

Eidenshink, J.C., and J.L. Faundeen, 1994. The 1 km AVHRR global land data set: First stages of implementation, *International Journal of Remote Sensing*, 15(17):3443–3462.

Gasch, J. and K.A. Campana, 2000. Cloud cover avoidance in space-based remote sensing acquisition, *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery IV* (Sylvia S. Shen and Michael R. Descour, editors), *Proceedings of SPIE*, Vol. 4049, pp. 336–347.

Gao, B.-C., and Y.J. Kaufman, 1995. Selection of the 1.375 μm MODIS channel for remote sensing of cirrus clouds and stratospheric aerosols from space, *Journal of Atmospheric Science*, 52:4231–4237.

Gao, B.-C., Y.J. Kaufman, W. Han, and W.J. Wiscombe, 1998. Correction of thin cirrus path radiance in the 0.4–1.0 μm spectral region using the sensitive 1.375 μm cirrus detecting channel, *Journal of Geophysical Research*, 103:32169–32176.

Goward, S.N., J. Haskett, D.L. Williams, T. Arvidson, J. Gasch, R. Lonigro, M. Reeley, J. Irons, R. Dubayah, S. Turner, K. Campana, R. Bindschadler, 1999. Enhanced Landsat capturing all the Earth's land areas, *EOS Transactions*, 80(26):289–293.

Irish, R., 1999. Landsat 7 Science Data Users Handbook, Landsat Project Science Office, Goddard Space Flight Center, URL: http://landsathandbook.gsfc.nasa.gov/handbook.html (last date accessed: 22).

Irish, R., 2000. Landsat 7 automatic cloud cover assessment: Algorithms for multispectral, hyperspectral, and ultraspectral imagery, *Proceedings of SPIE*, Vol. 4049, pp. 348–355.

Markham, B.L., and J.L. Barker, 1986. Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures, *EOSAT Landsat Technical Notes*, No. 1.

Markham, B.L., T.J. Arvidson, S.N. Goward, J. Barsi, and P. Scaramuzza, 2006. Landsat-7 long-term acquisition plan radiometry: Evolution over time, *Photogrammetric Engineering & Remote Sensing*, 72(10).

Platnick, S., M.D. King, S.A. Ackerman, W.P. Menzel, B.A. Baum, J.C. Riedi, and R.A. Frey, 2003. The MODIS cloud products: Algorithms and examples from Terra, *IEEE Transactions on Geoscience and Remote Sensing*, 41:459–473.